

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁶ : C07K 14/71, 14/62, 14/475, C12P 21/02, C12N 1/21, 5/10, 15/09</p>	<p>A1</p>	<p>(11) International Publication Number: WO 98/01475 (43) International Publication Date: 15 January 1998 (15.01.98)</p>
<p>(21) International Application Number: PCT/US96/11452 (22) International Filing Date: 9 July 1996 (09.07.96) (71) Applicants (for all designated States except US): SMITHK- LINE BEECHAM CORPORATION [US/US]; Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US). PRESIDENT AND FELLOWS OF HARVARD COLLEGE [US/US]; 17 Quincy Street, Cambridge, MA 02139 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): DUNNINGTON, Damien, J. [US/US]; 23 Forsythia Court, Lafayette Hills, PA 19444 (US). FRANTZ, James, D. [IS/US]; 22 Jenison Street, New- ton, MA 02169 (US). SHOELSON, Steven, E. [US/US]; 8 Mocassin Path, Natick, MA 01760 (US). (74) Agents: BAUMEISTER, Kirk et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).</p>		<p>(81) Designated States: AL, AM, AU, BB, BG, BR, CA, CN, CZ, EE, GE, HU, IL, IS, JP, KG, KP, KR, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published With international search report.</p>
<p>(54) Title: GROWTH FACTOR RECEPTOR-BINDING INSULIN RECEPTOR (57) Abstract Isolated nucleic acid encoding a growth factor receptor binding protein-insulin receptor isoform, protein obtainable from the nucleic acid, recombinant host cells transformed with the nucleic acid and use of the protein and nucleic acid sequence are disclosed.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LJ	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

GROWTH FACTOR RECEPTOR-BINDING INSULIN RECEPTOR**Field of the Invention**

The present invention relates to an isolated
5 isoform of human growth factor receptor-binding insulin-
receptor protein (GrbIR-1) gene; to essentially pure
human GrbIR-1 protein; and to compositions and methods
of producing and using human GrbIR-1 sequences and
proteins.

10

Background of the Invention

A number of polypeptide growth factors and hormones
mediate their cellular effects through a signal
transduction pathway. Transduction of signals from the
15 cell surface receptors for these ligands to
intracellular effectors frequently involves
phosphorylation or dephosphorylation of specific protein
substrates by regulatory protein tyrosine kinases (PTK)
and phosphatases. Tyrosine phosphorylation is a major
20 mediator of signal transduction in multicellular
organisms. Receptor-bound, membrane-bound and
intracellular PTKs regulate cell proliferation, cell
differentiation and signalling processes in immune
system cells.

25 Aberrant PTK activity has been implicated or is
suspected in a number of pathologies such as diabetes,
atherosclerosis, psoriasis, septic shock, bone loss,
anemia, many cancers and other proliferative diseases.
Accordingly, tyrosine kinases and the signal
30 transduction pathways which they are part of are
potential targets for drug design. For a review, see
Levitzki et al. in *Science* 267, 1782-1788 (1995).

Many of the proteins comprising signal transduction
pathways are present at low levels and often have
35 opposing activities. The properties of these signalling
molecules allow the cell to control transduction by
means of the subcellular location and juxtaposition of

effectors as well as by balancing activation with repression such that a small change in one pathway can achieve a switching effect.

The formation of transducing complexes by
5 juxtaposition of the signalling molecules through protein-protein interactions are mediated by specific docking domain sequence motifs. Src homology 2 (SH2) domains, which are conserved non-catalytic sequences of approximately 100 amino acids found in a variety of
10 signalling molecules such as non-receptor PTKs and kinase target effector molecules and in oncogenic proteins, play a critical role. The SH2 domains are highly specific for short phosphotyrosine-containing peptide sequences found in autophosphorylated PTK
15 receptors or intracellular tyrosine kinases. Src homology 3 (SH3) domains, conserved sequences of approximately 50 amino acids that mediate protein-protein interactions through sequence-specific binding to proline-rich motifs in target proteins, are also
20 critically involved in signal transduction. Pleckstrin homology (PH) domains are also involved in signal transduction and control membrane association of signaling molecules. See G. Shaw, *Bioessays* 18, 35-46 (1996). At least 90 proteins having conserved SH2, SH3
25 or PH domains, and, in many cases, distinct catalytic domains, are now known.

One approach towards the pharmacological regulation of signal transduction pathways is to design ligands which selectively bind to a chosen PH domain and thus
30 affect the interaction of membrane-associated inositol 1,4,5-trisphosphate with its PH domain-containing target molecule, thereby modulating signal transduction. Any selective modulators would provide a useful lead for drug development.

35 Growth factor receptor binding protein-Inuslin Receptor (Grb-IR) is a cytoplasmic signalling molecule containing an SH2 domain and a partial PH domain with a

pleckstrin homology domain of phospholipase C-delta 1 binds with high affinity to phosphatidylinositol 4,5-bisphosphate in bilayer membranes). However, the known grb-IR sequence lacks an intact PH domain.

5 The involvement of Grb-IR in the signal transduction of the insulin receptor pathway necessitates the identification of other human Grb-IR homologs and isoforms, preferably those containing intact PH domains, and their cDNAs. A need also exists
10 for compounds which modulate the activity of Grb-IR homologs and isoforms, for methods to identify such modulators and for reagents useful in such methods.

Summary of the Invention

15 Accordingly, one aspect of the present invention is an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding human GrbIR-1 having the nucleotide sequence as set forth in SEQ ID NO:1 from
20 nucleotide 289 to 1897;

(b) a polynucleotide capable of hybridizing to the complement of a polynucleotide according to (a) under moderately stringent hybridization conditions and which encodes a functional human GrbIR-1; and

25 (c) a degenerate polynucleotide according to (a) or (b).

Another aspect of the invention is a functional polypeptide encoded by the polynucleotides of the invention.

30 Another aspect of the invention is a method for preparing essentially pure human GrbIR-1 protein comprising culturing a recombinant host cell comprising a vector comprising a polynucleotide of the invention under conditions promoting expression of the protein and
35 recovery thereof.

wide tissue and cell distribution. The molecule was first described by F. Liu and R. A. Roth in *Proc Natl. Acad. Sci. USA* 92, 10287-10291 (1995). Interaction of Grb-IR with growth factor receptors such as the insulin receptor (IR) is mediated by the SH2 domain, can be dependent upon receptor tyrosine autophosphorylation and involves a direct interaction between Grb-IR and the phosphorylated receptors.

Further, binding of Grb-IR to the insulin receptor has been shown to inhibit subsequent signalling events such as insulin-dependent tyrosine phosphorylation of a 60k GAP-associated protein, IRS-1 and insulin induced association of phosphatidyl inositol-3 kinase with IRS-1 (Liu and Roth, *supra*). Thus, Grb-IR inhibits insulin signalling through the IR. Membrane association of signalling molecules is important for bringing them in close proximity to other effectors. An example is ras which is farnesylated at the C-terminus and thereby located to the plasma membrane. The importance of such localization is shown by the inhibitory effect of farnesyl transferase inhibitors on ras-mediated signal transduction. See Tamanoi, F., *Trends in Biochemical Sciences* 18, 349-353 (1993).

In the case of grb-IR, a PH domain could serve a similar purpose, since PH domains are known to facilitate membrane association of proteins through binding to inositol 1,4,5-trisphosphate residues in cell membranes. See H. F. Paterson et al., *Biochem. J.* 312, 661-666 (1995). Phospholipase C delta 1 requires a pleckstrin homology domain for interaction with the plasma membrane. See D. S. Wang & G. Shaw, *Biochem. Biophys. Res. Commun.* 217, 608-615 (1995). The association of the C-terminal region of beta I sigma II spectrin to brain membranes is mediated by a PH domain, does not require membrane proteins, and coincides with a inositol-1,4,5 triphosphate binding site. See P. Garcia et al., *Biochemistry* 34, 16228-16234 (1995). The

Another aspect of the invention is an antisense oligonucleotide comprising a sequence which is capable of binding to the polynucleotide of the invention.

Another aspect of the invention is a modulator of the polypeptides of the invention.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that modulates GrbIR-1 activity by affecting the binding of GrbIR-1 to cellular binding partners comprising the steps of:

- (a) providing a GrbIR-1 protein having the amino acid sequence of GrbIR-1 (SEQ ID NO:2) or a functional derivative thereof and a cellular binding partner or synthetic analog thereof;
- (b) incubating with a test substance which is suspected of modulating GrbIR-1 activity under conditions which permit the formation of a GrbIR-1 protein/cellular binding partner complex;
- (c) assaying for the presence of the complex, free GrbIR-1 protein or free cellular binding partner; and
- (d) comparing to a control to determine the effect of the substance.

Another aspect of the invention is a method for assaying for the presence of a substance that modulates GrbIR-1 activity by direct binding to GrbIR-1 protein comprising the steps of:

- (a) providing a labelled GrbIR-1 protein having the amino acid sequence of GrbIR-1 (SEQ ID NO:2) or a functional derivative thereof
- (b) providing solid support-associated modulator candidates;
- (c) incubating a mixture of the labelled GrbIR-1 protein with the support-associated modulator candidates under conditions which can permit the formation of a GrbIR-1 protein/modulator candidate complex;
- (d) separating the solid support from free soluble labelled GrbIR-1 protein;

(e) assaying for the presence of solid support-associated labelled protein;

(f) isolating the solid support complexed with labelled GrbIR-1 protein; and

5 (g) identifying the modulator candidate.

Another aspect of the invention is GrbIR-1 protein modulating compounds identified by the methods of the invention.

10 Another aspect of the invention is a method for the treatment of a patient having need to modulate GrbIR-1 activity comprising administering to the patient a therapeutically effective amount of the modulating compounds of the invention.

Another aspect of the invention is a method of treating conditions which are related to insufficient GrbIR-1 protein function which comprises:

(a) isolating cells from a patient deficient in GrbIR-1 protein function;

(b) altering the cells by transfecting the polynucleotide of claim 1 into the cells wherein a GrbIR-1 protein is expressed; and

(c) introducing the cells back to the patient to alleviate the condition.

Another aspect of the invention is a method of treating conditions which are related to insufficient GrbIR-1 protein function which comprises administering the polynucleotide of claim 1 to a patient deficient in GrbIR-1 protein function wherein a GrbIR-1 protein is expressed and alleviates the condition.

Another aspect of the invention is a transgenic non-human animal capable of expressing in any cell thereof the DNA encoding the polypeptides of the invention.

Brief Description of the Drawings

Figure 1 is a multiple amino acid sequence alignment of GrbIR-1, Grb-IR, murine Grb10 and human Grb7.

- 5 Figure 2 is an amino acid sequence alignment of human GrbIR-1 with human Grb-IR.

Detailed Description of the Invention

- As used herein, the term "GrbIR-1 gene" refers to
- 10 DNA molecules comprising a nucleotide sequence that encodes an isoform of human growth factor receptor binding insulin receptor. The GrbIR-1 gene sequence is listed in SEQ ID NO:1. The coding region of the GrbIR-1 gene consists of nucleotides 289 to 1897 of SEQ ID NO:1.
- 15 The deduced 536 amino acid sequence of the GrbIR-1 gene product GrbIR-1 is listed in SEQ ID NO:2.

- As used herein, the term "functional fragments" when used to modify a specific gene or gene product means a less than full length portion of the gene or
- 20 gene product which retains substantially all of the biological function associated with the full length gene or gene product to which it relates. To determine whether a fragment of a particular gene or gene product is a functional fragment, fragments are generated by
- 25 well-known nucleolytic or proteolytic techniques or by the polymerase chain reaction and the fragments tested for the described biological function.

- As used herein, an "antigen" refers to a molecule containing one or more epitopes that will stimulate a
- 30 host's immune system to make a humoral and/or cellular antigen-specific response. The term is also used herein interchangeably with "immunogen."

- As used herein, the term "epitope" refers to the site on an antigen or hapten to which a specific
- 35 antibody molecule binds. The term is also used herein interchangeably with "antigenic determinant" or "antigenic determinant site."

As used herein, "monoclonal antibody" is understood to include antibodies derived from one species (e.g., murine, rabbit, goat, rat, human, etc.) as well as antibodies derived from two (or perhaps more) species (e.g., chimeric and humanized antibodies).

As used herein, a coding sequence is "operably linked to" another coding sequence when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one another so long as the expressed sequence is ultimately processed to produce the desired protein.

As used herein, "recombinant" polypeptides refer to polypeptides produced by recombinant DNA techniques; i.e., produced from cells transformed by an exogenous DNA construct encoding the desired polypeptide. "Synthetic" polypeptides are those prepared by chemical synthesis.

As used herein, a "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication *in vivo*; i.e., capable of replication under its own control.

As used herein, a "vector" is a replicon, such as a plasmid, phage, or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

As used herein, a "reference" gene refers to the wild type human GrbIR-1 gene sequence of the invention and is understood to include the various sequence polymorphisms that exist, wherein nucleotide substitutions in the gene sequence exist, but do not affect the essential function of the gene product.

As used herein, a "mutant" gene refers human GrbIR-1 sequences different from the reference gene wherein nucleotide substitutions and/or deletions and/or insertions result in perturbation of the essential function of the gene product.

As used herein, a DNA "coding sequence of" or a "nucleotide sequence encoding" a particular protein, is a DNA sequence which is transcribed and translated into a polypeptide when placed under the control of appropriate regulatory sequences.

As used herein, a "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bound at its 3' terminus by a translation start codon (e.g., ATG) of a coding sequence and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes. Prokaryotic promoters contain Shine-Dalgarno sequences in addition to the -10 and -35 consensus sequences.

As used herein, DNA "control sequences" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers and the like, which collectively provide for the expression (i.e., the transcription and translation) of a coding sequence in a host cell.

As used herein, a control sequence "directs the expression" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

As used herein, a "host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous DNA sequence.

5 As used herein, a cell has been "transformed" by exogenous DNA when such exogenous DNA has been introduced inside the cell membrane. Exogenous DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In
10 prokaryotes and yeasts, for example, the exogenous DNA may be maintained on an episomal element, such as a plasmid. With respect to eukaryotic cells, a stably transformed or transfected cell is one in which the exogenous DNA has become integrated into the chromosome
15 so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the exogenous DNA.

20 As used herein, "transfection" or "transfected" refers to a process by which cells take up foreign DNA and integrate that foreign DNA into their chromosome. Transfection can be accomplished, for example, by various techniques in which cells take up DNA (e.g.,
25 calcium phosphate precipitation, electroporation, assimilation of liposomes, etc.) or by infection, in which viruses are used to transfer DNA into cells.

As used herein, a "target cell" is a cell that is selectively transfected over other cell types (or cell
30 lines).

As used herein, a "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many
35 generations.

As used herein, a "heterologous" region of a DNA construct is an identifiable segment of DNA within or attached to another DNA molecule that is not found in association with the other molecule in nature. Thus, when the heterologous region encodes a gene, the gene will usually be flanked by DNA that does not flank the gene in the genome of the source animal. Another example of a heterologous coding sequence is a construct where the coding sequence itself is not found in nature (e.g., synthetic sequences having codons different from the native gene). Allelic variation or naturally occurring mutational events do not give rise to a heterologous region of DNA, as used herein.

As used herein, a "modulator" of a polypeptide is a substance which can affect the polypeptide function.

An aspect of the present invention is isolated polynucleotides encoding a human GrbIR-1 protein including substantially similar sequences and functional fragments. Isolated polynucleotide sequences are substantially similar if they are capable of hybridizing under moderately stringent conditions to SEQ ID NO:1 or they encode DNA sequences which are degenerate to SEQ ID NO:1 or are degenerate to those sequences capable of hybridizing under moderately stringent conditions to SEQ ID NO:1.

Moderately stringent conditions is a term understood by the skilled artisan and has been described in, for example, Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2nd edition, Vol. 1, pp. 101-104, Cold Spring Harbor Laboratory Press (1989). An exemplary hybridization protocol using moderately stringent conditions is as follows. Nitrocellulose filters are prehybridized at 65°C in a solution containing 6X SSPE, 5X Denhardt's solution (10g Ficoll, 10g BSA and 10g polyvinylpyrrolidone per liter solution), 0.05% SDS and 100 ug/ml tRNA. Hybridization

probes are labeled, preferably radiolabelled (e.g., using the Bios TAG-IT® kit). Hybridization is then carried out for approximately 18 hours at 65°C. The filters are then washed twice in a solution of 2X SSC and 0.5% SDS at room temperature for 15 minutes. Subsequently, the filters are washed at 58°C, air-dried and exposed to X-ray film overnight at -70°C with an intensifying screen.

Degenerate DNA sequences encode the same amino acid sequence as SEQ ID NO:2 or the proteins encoded by that sequence capable of hybridizing under moderately stringent conditions to SEQ ID NO:1, but have variation(s) in the nucleotide coding sequences because of the degeneracy of the genetic code. For example, the degenerate codons UUC and UUU both code for the amino acid phenylalanine, whereas the four codons GGX all code for glycine.

Alternatively, substantially similar sequences are defined as those sequences in which about 70%, preferably about 80% and most preferably about 90%, of the nucleotides or amino acids match over a defined length of the molecule. As used herein, substantially similar refers to the sequences having similar identity to the sequences of the instant invention. Thus nucleotide sequences that are substantially the same can be identified by hybridization or by sequence comparison. Protein sequences that are substantially the same can be identified by techniques such as proteolytic digestion, gel electrophoresis and/or microsequencing. Excluded from the definition of substantially similar sequences is Grb-IR.

Embodiments of the isolated polynucleotides of the invention include DNA, genomic DNA and RNA, preferably of human origin. A method for isolating a nucleic acid molecule encoding a GrbIR-1 protein is to probe a genomic or cDNA library with a natural or artificially

designed probe using art. recognized procedures. See, e.g., "Current Protocols in Molecular Biology", Ausubel et al. (eds.) Greene Publishing Association and John Wiley Interscience, New York, 1989, 1992. The ordinarily skilled artisan will appreciate that SEQ ID NO:1 or fragments thereof comprising at least 15 contiguous nucleotides are particularly useful probes. It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include, but are not limited to, radioisotopes, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes would enable the ordinarily skilled artisan to isolate complementary copies of genomic DNA, cDNA or RNA polynucleotides encoding GrbIR-1 proteins from human, mammalian or other animal sources or to screen such sources for related sequences, e.g., additional members of the family, type and/or subtype, including transcriptional regulatory and control elements as well as other stability, processing, translation and tissue specificity-determining regions from 5' and/or 3' regions relative to the coding sequences disclosed herein, all without undue experimentation.

Another aspect of the invention is functional polypeptides encoded by the polynucleotides of the invention. An embodiment of a functional polypeptide of the invention is the human GrbIR-1 protein having the amino acid sequence set forth in SEQ ID NO:2.

Another aspect of the invention is a method for preparing essentially pure human GrbIR-1 protein. Yet another aspect is the human GrbIR-1 protein produced by the preparation method of the invention. This protein has the amino acid sequence listed in SEQ ID NO:2 and includes variants with a substantially similar amino acid sequence that have the same function. The proteins of this invention are preferably made by recombinant genetic

engineering techniques by culturing a recombinant host cell containing a vector encoding the polynucleotides of the invention under conditions promoting the expression of the protein and recovery thereof.

5 The isolated polynucleotides, particularly the DNAs, can be introduced into expression vectors by operatively linking the DNA to the necessary expression control regions, e.g., regulatory regions, required for gene expression. The vectors can be introduced into an
10 appropriate host cell such as a prokaryotic, e.g., bacterial, or eukaryotic, e.g., yeast or mammalian cell by methods well known in the art. See Ausubel et al., *supra*. The coding sequences for the desired proteins, having been prepared or isolated, can be cloned into any
15 suitable vector or replicon. Numerous cloning vectors are known to those of skill in the art and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include, but are not
20 limited to, the bacteriophage λ (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pGEX4T-3 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and
25 *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), YIp5 (*Saccharomyces*), a baculovirus insect cell system, a *Drosophila* insect system, YCp19 (*Saccharomyces*) and pSV2neo (mammalian cells). See generally, "DNA Cloning":
30 Vols. I & II, Glover et al. ed. IRL Press Oxford (1985) (1987); and T. Maniatis et al. ("Molecular Cloning" Cold Spring Harbor Laboratory (1982).

 The gene can be placed under the control of control elements such as a promoter, ribosome binding site (for
35 bacterial expression) and, optionally, an operator, so that the DNA sequence encoding the desired protein is

transcribed into RNA in the host cell transformed by a vector containing the expression construct. The coding sequence may or may not contain a signal peptide or leader sequence. The proteins of the present invention
5 can be expressed using, for example, the *E. coli* tac promoter or the protein A gene (*spa*) promoter and signal sequence. Leader sequences can be removed by the bacterial host in post-translational processing. See, e.g., U.S. Patent Nos. 4,431,739; 4,425,437 and
10 4,338,397.

In addition to control sequences, it may be desirable to add regulatory sequences which allow for regulation of the expression of the protein sequences relative to the growth of the host cell. Regulatory
15 sequences are known to those of skill in the art. Exemplary are those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound or to various temperature or
20 metabolic conditions. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

An expression vector is constructed so that the particular coding sequence is located in the vector with
25 the appropriate regulatory sequences, the positioning and orientation of the coding sequence with respect to the control sequences being such that the coding sequence is transcribed under the "control" of the control sequences, i.e., RNA polymerase which binds to
30 the DNA molecule at the control sequences transcribes the coding sequence. Modification of the sequences encoding the particular antigen of interest may be desirable to achieve this end. For example, in some cases it may be necessary to modify the sequence so that
35 it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the reading frame. The control sequences and other regulatory

sequences may be ligated to the coding sequence prior to insertion into a vector, such as the cloning vectors described above. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

In some cases, it may be desirable to produce mutants or analogues of human GrbIR-1 protein. Mutants or analogues may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. See, e.g., T. Maniatis et al., *supra*; "DNA Cloning," Vols. I and II, *supra*; and "Nucleic Acid Hybridization", *supra*.

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described above under conditions whereby the protein of interest is expressed. Preferred mammalian cells include human embryonic kidney cells (293), monkey kidney cells, fibroblast (COS) cells, Chinese hamster ovary (CHO) cells, *Drosophila* or murine L-cells. If the expression system secretes the protein into growth media, the protein can be purified directly from the media. If the protein is not secreted, it is isolated from cell lysates or recovered from the cell membrane fraction. The selection of the appropriate growth conditions and recovery methods are within the skill of the art.

An alternative method to identify proteins of the present invention is by constructing gene libraries, using the resulting clones to transform *E. coli* and pooling and screening individual colonies using

polyclonal serum or monoclonal antibodies to human GrbIR-1.

The proteins of the present invention may also be produced by chemical synthesis such as solid phase peptide synthesis on an automated peptide synthesizer, using known amino acid sequences or amino acid sequences derived from the DNA sequence of the genes of interest. Such methods are known to those skilled in the art.

The proteins of the present invention or their fragments comprising at least one epitope can be used to produce antibodies, both polyclonal and monoclonal, directed to epitopes corresponding to amino acid sequences disclosed herein. If polyclonal antibodies are desired, a selected mammal such as a mouse, rabbit, goat or horse is immunized with a protein of the present invention, or its fragment, or a mutant protein. Serum from the immunized animal is collected and treated according to known procedures. Serum polyclonal antibodies can be purified by immunoaffinity chromatography or other known procedures.

Monoclonal antibodies to the proteins of the present invention, and to the fragments thereof, can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by using hybridoma technology is well known. Immortal antibody-producing cell lines can be created by cell fusion and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA or transfection with Epstein-Barr virus. See, e.g., M. Schreier et al., "Hybridoma Techniques" (1980); Hammerling et al., "Monoclonal Antibodies and T-cell Hybridomas" (1981); Kennett et al., "Monoclonal Antibodies" (1980); and U.S. Patent Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,452,570; 4,466,917; 4,472,500; 4,491,632; and 4,493,890. Panels of monoclonal antibodies produced against the antigen of interest, or fragment thereof, can be screened for

various properties, i.e., for isotype, epitope, affinity, etc. Monoclonal antibodies are useful in purification, using immunoaffinity techniques, of the individual antigens which they are directed against.

5 Alternatively, genes encoding the monoclonals of interest may be isolated from the hybridomas by PCR techniques known in the art and cloned and expressed in the appropriate vectors. The antibodies of this invention, whether polyclonal or monoclonal have

10 additional utility in that they may be employed as reagents in immunoassays, RIA, ELISA, and the like. The antibodies of the invention can be labeled with an analytically detectable reagent such as a radioisotope, fluorescent molecule or enzyme.

15 Chimeric antibodies, in which non-human variable regions are joined or fused to human constant regions (see, e.g., Liu et al., *Proc. Natl Acad. Sci. USA*, 84, 3439 (1987)), may also be used in assays or therapeutically. Preferably, a therapeutic monoclonal

20 antibody would be "humanized" as described in Jones et al., *Nature*, 321, 522 (1986); Verhoeyen et al., *Science*, 239, 1534 (1988); Kabat et al., *J. Immunol.*, 147, 1709 (1991); Queen et al., *Proc. Natl Acad. Sci. USA*, 86, 10029 (1989); Gorman et al., *Proc. Natl Acad. Sci. USA*, 88, 34181 (1991); and Hodgson et al., *Bio/Technology*, 9:, 421 (1991).

25

Another aspect of the present invention is modulators of the polypeptides of the invention. Functional modulation of GrbIR-1 by a substance includes

30 partial to complete inhibition of function, identical function, as well as enhancement of function. Embodiments of modulators of the invention include peptides, oligonucleotides and small organic molecules including peptidomimetics.

35 Another aspect of the invention is antisense oligonucleotides comprising a sequence which is capable of binding to the polynucleotides of the invention.

Synthetic oligonucleotides or related antisense chemical structural analogs can be designed to recognize, specifically bind to and prevent transcription of a target nucleic acid encoding GrbIR-1 protein by those of
5 ordinary skill in the art. See generally, Cohen, J.S., *Trends in Pharm. Sci.*, 10, 435(1989) and Weintraub, H.M., *Scientific American*, January (1990) at page 40.

Another aspect of the invention is a method for assaying a medium for the presence of a substance that
10 modulates GrbIR-1 protein function by affecting the binding of GrbIR-1 protein to cellular binding partners. Examples of modulators include, but are not limited to peptides and small organic molecules including peptidomimetics. A GrbIR-1 protein is provided having
15 the amino acid sequence of human GrbIR-1 (SEQ ID NO:2) or a functional derivative thereof together with a cellular binding partner or synthetic analog thereof. The mixture is incubated with a test substance which is suspected of modulating GrbIR-1 activity, under
20 conditions which permit the formation of a GrbIR-1 gene product/cellular binding partner complex. An assay is performed for the presence of the complex, free GrbIR-1 protein or free cellular binding partner and the result compared to a control to determine the effect of the
25 test substance.

Another aspect of the invention is a method for assaying for the presence of a substance that modulates GrbIR-1 activity by direct binding to GrbIR-1 protein. Examples of modulators include, but are not limited to,
30 peptides and small organic molecules including peptidomimetics. Modulator candidates are synthesized on a solid support by techniques such as those disclosed in Lam et al., *Nature* 354, 82 (1991) or Burbaum et al., *Proc. Natl. Acad. Sci. USA* 92, 6027 (1995) to provide
35 solid support-associated modulator candidates. A labelled GrbIR-1 protein is provided having the amino acid sequence of human GrbIR-1 (SEQ ID NO:2) or a

functional derivative thereof. Exemplary labels include directly attached fluorescent or colored dyes, biotin, radioisotopes or epitope tags, which are detectable by a suitable antibody. A mixture of solid support-associated modulator candidates and labelled GrbIR-1 protein is incubated under conditions which can permit the formation of a GrbIR-1 protein/modulator candidate complex. The solid support is separated from free soluble labelled GrbIR-1 protein. An assay is performed for the presence of solid support-associated labelled protein. Solid supports complexed with labelled protein are isolated and the identity of the modulator candidate determined by techniques well known to those skilled in the art.

Modulation of GrbIR-1 function would be expected to be useful for treatment of diabetes. Inhibition of grbIR-1 could be effected through antagonism of the SH2 domain/phosphorylated IR interaction or through inhibition of the binding of the PH domain to phosphatidylinositol 4,5-bisphosphate.

Further, GrbIR-1 could be used to isolate proteins which interact with it and this interaction could be a target for interference. Inhibitors of protein-protein interactions between GrbIR-1 and other factors could lead to the development of pharmaceutical agents for the modulation of GrbIR-1 activity.

Methods to assay for protein-protein interactions, such as that of a GrbIR-1 gene product/binding partner complex, and to isolate proteins interacting with GrbIR-1 are known to those skilled in the art. Use of the methods discussed below enable one of ordinary skill in the art to accomplish these aims without undue experimentation.

The yeast two-hybrid system provides methods for detecting the interaction between a first test protein and a second test protein, *in vivo*, using reconstitution of the activity of a transcriptional activator. The

method is disclosed in U.S. Patent No. 5,283,173;

reagents are available from Clontech and Stratagene.

Briefly, GrbIR-1 cDNA is fused to a Gal4 transcription factor DNA binding domain and expressed in yeast cells.

- 5 cDNA library members obtained from cells of interest are fused to a transactivation domain of Gal4. cDNA clones which express proteins which can interact with GrbIR-1 will lead to reconstitution of Gal4 activity and transactivation of expression of a reporter gene such as
- 10 *Gall-lacZ*. Optionally, the host cells can be co-transfected with a protein tyrosine kinase to induce tyrosine phosphorylation of members of the cDNA library. Such phosphorylation is necessary for optimum interaction with the SH2 domain of GrbIR-1.

- 15 An alternative method is screening of λ gt11, λ ZAP (Stratagene) or equivalent cDNA expression libraries with recombinant GrbIR-1. Recombinant GrbIR-1 protein or fragments thereof are fused to small peptide tags such as FLAG, HSV or GST. The peptide tags can possess
- 20 convenient phosphorylation sites for a kinase such as heart muscle creatine kinase or they can be biotinylated. Recombinant GrbIR-1 can be phosphorylated with 32 [P] or used unlabeled and detected with streptavidin or antibodies against the tags. λ gt11cDNA
- 25 expression libraries are made from cells of interest and are incubated with the recombinant GrbIR-1, washed and cDNA clones isolated which interact with GrbIR-1. See, e.g., T. Maniatis et al, *supra*.

- Another method is the screening of a mammalian
- 30 expression library in which the cDNAs are cloned into a vector between a mammalian promoter and polyadenylation site and transiently transfected in COS or 293 cells followed by detection of the binding protein 48 hours later by incubation of fixed and washed cells with a
- 35 labelled GrbIR-1, preferably iodinated, and detection of bound GrbIR-1 by autoradiography. See Sims et al.,

Science 241, 585-589 (1988) and McMahon et al., *EMBO J.* 10, 2821-2832 (1991). In this manner, pools of cDNAs containing the cDNA encoding the binding protein of interest can be selected and the cDNA of interest can be isolated by further subdivision of each pool followed by cycles of transient transfection, binding and autoradiography. Alternatively, the cDNA of interest can be isolated by transfecting the entire cDNA library into mammalian cells and panning the cells on a dish containing GrbIR-1 bound to the plate. Cells which attach after washing are lysed and the plasmid DNA isolated, amplified in bacteria, and the cycle of transfection and panning repeated until a single cDNA clone is obtained. See Seed et al., *Proc. Natl. Acad. Sci. USA* 84, 3365 (1987) and Aruffo et al., *EMBO J.* 6, 3313 (1987). If the binding protein is secreted, its cDNA can be obtained by a similar pooling strategy once a binding or neutralizing assay has been established for assaying supernatants from transiently transfected cells. General methods for screening supernatants are disclosed in Wong et al., *Science* 228, 810-815 (1985).

Another alternative method is isolation of proteins interacting with GrbIR-1 directly from cells. Fusion proteins of GrbIR-1 with GST or small peptide tags are made and immobilized on beads. Biosynthetically labeled or unlabeled protein extracts from the cells of interest are prepared, incubated with the beads and washed with buffer. Proteins interacting with GrbIR-1 are eluted specifically from the beads and analyzed by SDS-PAGE. Binding partner primary amino acid sequence data are obtained by microsequencing. Optionally, the cells can be treated with agents that induce a functional response such as tyrosine phosphorylation of cellular proteins. An example of such an agent would be a growth factor or cytokine such as interleukin-2.

Another alternative method is immunoaffinity purification. Recombinant GrbIR-1 is incubated with

labeled or unlabeled cell extracts and immunoprecipitated with anti-GrbIR-1 antibodies. The immunoprecipitate is recovered with protein A-Sepharose and analyzed by SDS-PAGE. Unlabelled proteins are
5 labeled by biotinylation and detected on SDS gels with streptavidin. Binding partner proteins are analyzed by microsequencing. Further, standard biochemical purification steps known to those skilled in the art may be used prior to microsequencing.

10 Yet another alternative method is screening of peptide libraries for binding partners. Recombinant tagged or labeled GrbIR-1 is used to select peptides from a peptide or phosphopeptide library which interact with GrbIR-1. Sequencing of the peptides leads to
15 identification of consensus peptide sequences which might be found in interacting proteins.

GrbIR-1 binding partners identified by any of these methods or other methods which would be known to those of ordinary skill in the art as well as those putative
20 binding partners discussed above can be used in the assay method of the invention. Assaying for the presence of GrbIR-1/binding partner complex are accomplished by, for example, the yeast two-hybrid system, ELISA or immunoassays using antibodies specific
25 for the complex. In the presence of test substances which interrupt or inhibit formation of GrbIR-1/binding partner interaction, a decreased amount of complex will be determined relative to a control lacking the test substance.

30 Assays for free GrbIR-1 or binding partner are accomplished by, for example, ELISA or immunoassay using specific antibodies or by incubation of radiolabeled GrbIR-1 with cells or cell membranes followed by centrifugation or filter separation steps. In the
35 presence of test substances which interrupt or inhibit formation of GrbIR-1/binding partner interaction, an increased amount of free GrbIR-1 or free binding partner

will be determined relative to a control lacking the test substance.

Another aspect of the invention is pharmaceutical compositions comprising an effective amount of a GrbIR-1 modulator of the invention and a pharmaceutically acceptable carrier. Pharmaceutical compositions of modulators of this invention for parenteral administration, i.e., subcutaneously, intramuscularly or intravenously or oral administration can be prepared.

The compositions for parenteral administration will commonly comprise a solution of the modulators of the invention or a cocktail thereof dissolved in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be employed, e.g., water, buffered water, 0.4% saline, 0.3% glycine and the like. These solutions are sterile and generally free of particulate matter. These solutions may be sterilized by conventional, well-known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, etc. The concentration of the modulator of the invention in such pharmaceutical formulation can vary widely, i.e., from less than about 0.5%, usually at or at least about 1% to as much as 15 or 20% by weight and will be selected primarily based on fluid volumes, viscosities, etc. according to the particular mode of administration selected.

Thus, a pharmaceutical composition of the modulator of the invention for intramuscular injection could be prepared to contain 1 mL sterile buffered water, and 50 mg of a protein of the invention. Similarly, a pharmaceutical composition of the modulator of the invention for intravenous infusion could be made up to contain 250 ml of sterile Ringer's solution, and 150 mg of a modulator of the invention. Actual methods for preparing parenterally administrable compositions are

well known or will be apparent to those skilled in the art and are described in more detail in, for example, *Remington's Pharmaceutical Science*, 15th ed., Mack Publishing Company, Easton, Pennsylvania.

- 5 The physician will determine the dosage of the present therapeutic agents which will be most suitable and it will vary with the form of administration and the particular compound chosen, and furthermore, it will vary with the particular patient under treatment.
- 10 Generally, the physician will wish to initiate treatment with small dosages substantially less than the optimum dose of the compound and increase the dosage by small increments until the optimum effect under the circumstances is reached. It will generally be found
- 15 that when the composition is administered orally, larger quantities of the active agent will be required to produce the same effect as a smaller quantity given parenterally. The therapeutic dosage will generally be from 1 to 10 milligrams per day and higher although it
- 20 may be administered in several different dosage units.

- Depending on the patient condition, the pharmaceutical composition of the invention can be administered for prophylactic and/or therapeutic treatments. In therapeutic application, compositions
- 25 are administered to a patient already suffering from a disease in an amount sufficient to cure or at least partially arrest the disease and its complications. In prophylactic applications, compositions containing the present compounds or a cocktail thereof are administered
- 30 to a patient not already in a disease state to enhance the patient's resistance to the disease.

- Single or multiple administrations of the pharmaceutical compositions can be carried out with dose levels and pattern being selected by the treating
- 35 physician. In any event, the pharmaceutical composition of the invention should provide a quantity of the

modulators of the invention sufficient to effectively treat the patient.

Additionally, some diseases result from inherited defective genes. These genes can be detected by
5 comparing the sequence of the defective gene with that of a normal one. Individuals carrying mutations in the GrbIR-1 gene may be detected at the DNA level by a variety of techniques. Nucleic acids used for diagnosis (genomic DNA, mRNA, etc.) may be obtained from a
10 patient's cells, such as from blood, urine, saliva or tissue biopsy, e.g., chorionic villi sampling or removal of amniotic fluid cells and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR, ligase chain
15 reaction (LCR), strand displacement amplification (SDA), etc. prior to analysis. See, e.g., Saiki et al., *Nature*, 324, 163-166 (1986), Bej, et al., *Crit. Rev. Biochem. Molec. Biol.*, 26, 301-334 (1991), Birkenmeyer et al., *J. Virol. Meth.*, 35, 117-126 (1991), Van Brunt, J., *Bio/Technology*, 8, 291-294 (1990)). RNA or cDNA may also be used for the same purpose. As an example, PCR
20 primers complementary to the nucleic acid of the instant invention can be used to identify and analyze GrbIR-1 mutations. For example, deletions and insertions can be
25 detected by a change in size of the amplified product in comparison to the normal GrbIR-1 genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled GrbIR-1 RNA of the invention or alternatively, radiolabelled GrbIR-1 antisense DNA
30 sequences of the invention. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures (T_m). Such a diagnostic would be particularly useful for prenatal and even neonatal testing.

35 In addition, point mutations and other sequence differences between the reference gene and "mutant" genes can be identified by yet other well-known

techniques, e.g., direct DNA sequencing, single-strand conformational polymorphism. See Orita et al., *Genomics*, 5, 874-879 (1989). For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotides or by automatic sequencing procedures with fluorescent-tags. Cloned DNA segments may also be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. The presence of nucleotide repeats may correlate to a causative change in GrbIR-1 activity or serve as marker for various polymorphisms.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures. See, e.g., Myers et al., *Science*, 230, 1242 (1985). In addition, sequence alterations, in particular small deletions, may be detected as changes in the migration pattern of DNA heteroduplexes in non-denaturing gel electrophoresis such as heteroduplex electrophoresis. See, e.g., Nagamine et al., *Am. J. Hum. Genet.*, 45, 337-339 (1989). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method as disclosed by Cotton et al. in *Proc. Natl. Acad. Sci. USA*, 85, 4397-4401 (1985).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization (e.g.,

heteroduplex electroporation, see, White et al.,
Genomics, 12, 301-306 (1992), RNase protection (e.g.,
Myers et al., Science, 230, 1242 (1985)) chemical
cleavage (e.g., Cotton et al., Proc. Natl. Acad. Sci.
5 USA, 85, 4397-4401 (1985))), direct DNA sequencing, or
the use of restriction enzymes (e.g., restriction
fragment length polymorphisms (RFLP) in which variations
in the number and size of restriction fragments can
indicate insertions, deletions, presence of nucleotide
10 repeats and any other mutation which creates or destroys
an endonuclease restriction sequence). Southern blotting
of genomic DNA may also be used to identify large (i.e.,
greater than 100 base pair) deletions and insertions.

In addition to conventional gel electrophoresis and
15 DNA sequencing, mutations such as microdeletions,
aneuploidies, translocations, inversions, can also be
detected by *in situ* analysis. See, e.g., Keller et al.,
DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA
(1993). That is, DNA or RNA sequences in cells can be
20 analyzed for mutations without isolation and/or
immobilization onto a membrane. Fluorescence *in situ*
hybridization (FISH) is presently the most commonly
applied method and numerous reviews of FISH have
appeared. See, e.g., Trachuck et al., Science, 250,
25 559-562 (1990), and Trask et al., Trends, Genet., 7,
149-154 (1991). Hence, by using nucleic acids based on
the structure of the GrbIR-1 genes, one can develop
diagnostic tests for genetic mutations.

In addition, some diseases are a result of, or are
30 characterized by, changes in gene expression which can
be detected by changes in the mRNA. Alternatively, the
GrbIR-1 gene can be used as a reference to identify
individuals expressing an increased or decreased level
of GrbIR-1 protein, e.g., by Northern blotting or *in*
35 *situ* hybridization.

Defining appropriate hybridization conditions is
within the skill of the art. See, e.g., "Current

Protocols in Mol. Biol." Vol. I & II, Wiley Interscience. Ausbel et al. (eds.) (1992). Probing technology is well known in the art and it is appreciated that the size of the probes can vary widely but it is preferred that the probe be at least 15 nucleotides in length. It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioisotopes, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. As a general rule, the more stringent the hybridization conditions the more closely related genes will be that are recovered.

The putative role of GrbIR-1 in signal transduction of the insulin receptor pathway establishes yet another aspect of the invention which is gene therapy. "Gene therapy" means gene supplementation where an additional reference copy of a gene of interest is inserted into a patient's cells. As a result, the protein encoded by the reference gene corrects the defect and permits the cells to function normally, thus alleviating disease symptoms. The reference copy would be a wild-type form of the GrbIR-1 gene or a gene encoding a protein or peptide which modulates the activity of the endogenous GrbIR-1.

Gene therapy of the present invention can occur *in vivo* or *ex vivo*. *Ex vivo* gene therapy requires the isolation and purification of patient cells, the introduction of a therapeutic gene and introduction of the genetically altered cells back into the patient. A replication-deficient virus such as a modified retrovirus can be used to introduce the therapeutic GrbIR-1 gene into such cells. For example, mouse Moloney leukemia virus (MMLV) is a well-known vector in clinical gene therapy trials. See, e.g., Boris-Lauerie et al., *Curr. Opin. Genet. Dev.*, 3, 102-109 (1993).

In contrast, *in vivo* gene therapy does not require isolation and purification of a patient's cells. The therapeutic gene is typically "packaged" for administration to a patient such as in liposomes or in a replication-deficient virus such as adenovirus as described by Berkner, K.L., in *Curr. Top. Microbiol. Immunol.*, 158, 39-66 (1992) or adeno-associated virus (AAV) vectors as described by Muzyczka, N., in *Curr. Top. Microbiol. Immunol.*, 158, 97-129 (1992) and U.S. Patent No. 5,252,479. Another approach is administration of "naked DNA" in which the therapeutic gene is directly injected into the bloodstream or muscle tissue. Another approach is administration of "naked DNA" in which the therapeutic gene is introduced into the target tissue by microparticle bombardment using gold particles coated with the DNA.

Cell types useful for gene therapy of the present invention include lymphocytes, hepatocytes, myoblasts, fibroblasts, any cell of the eye such as retinal cells, epithelial and endothelial cells. Preferably the cells are T lymphocytes drawn from the patient to be treated, hepatocytes, any cell of the eye or respiratory or pulmonary epithelial cells. Transfection of pulmonary epithelial cells can occur via inhalation of a nebulized preparation of DNA vectors in liposomes, DNA-protein complexes or replication-deficient adenoviruses. See, e.g., U.S. Patent No. 5,240,846.

Another aspect of the invention is transgenic, non-human mammals capable of expressing the polynucleotides of the invention in any cell. Transgenic, non-human animals may be obtained by transfecting appropriate fertilized eggs or embryos of a host with the polynucleotides of the invention or with mutant forms found in human diseases. See, e.g., U.S. Patent Nos. 4,736,866; 5,175,385; 5,175,384 and 5,175,386. The resultant transgenic animal may be used as a model for the study of *GrbIR-1* gene function or for producing

large amounts of grb-IR-1 protein for screening or crystallography purposes. Particularly useful transgenic animals are those which display a detectable phenotype associated with the expression of the GrbIR-1 protein. Drug development candidates may then be screened for their ability to reverse or exacerbate the relevant phenotype.

The present invention will now be described with reference to the following specific, non-limiting examples.

Example 1

GrbIR-1 full-length cDNA Cloning and Sequence Analysis

A search of a random cDNA sequence database consisting of short partial sequences known as expressed sequence tags (ESTs) with SH2 domain encoding sequences using the BLASTX algorithm disclosed an EST which was homologous to a murine epidermal growth factor receptor-binding protein *grb7* cDNA sequence reported by Margolis, B.L. et al. in *Proc. Natl. Acad. Sci. USA* 89, 8894-8898 (1992) (SEQ ID NO: 3). The EST was originally isolated from a human cerebellum cDNA library.

A 5'-rapid amplification of cDNA ends (5' RACE) protocol was used to isolate the 5' cDNA end of the putative human gene. Candidate 5' RACE products were amplified by PCR from a λ gt11 human skeletal muscle library (Clontech cat no. HL1124b). The PCR contained 100 ng of phage DNA, a lambda-specific primer 5'GATTGGTGGCGACGACTCC3' (SEQ ID NO: 4) and a gene-specific primer 5'CCCGTGAAACCAGTGCTGTG3' (SEQ ID NO: 5). Thirty cycles were conducted as follows: 94°C for 20 s, 70°C to 55°C in 0.5°C increments/cycle for 30 s and 72°C for 2 min. A PCR product of 1.7 kb was purified and subcloned into pBluescript II and sequenced. Sequence analysis revealed the fragment to be the 5' end of the

gene, containing the remaining coding sequence, including the N-terminus.

A cDNA encoding an intact coding sequence was assembled. A 3.4 kb PCR product was amplified from the EST using the primers T7 5'GTAATACGACTCACTATAGGGC3' (SEQ ID NO: 6) and 5'GGTAGCCAAAGTCCCCTCCA3' (SEQ ID NO: 7), and a 1.7 kb PCR product was amplified from the 5' RACE fragment isolated above using the primers 5'GATTGGTGGCGACGACTCC3' (SEQ ID NO: 8) and 5'TGGAGGGGACTTTGGCTACC3' (SEQ ID NO: 9). The PCR conditions were 94°C for 15 s, 55°C for 20 s, 72°C for 4 min., for 25 cycles. These products were combined by PCR in a second reaction containing each of the above PCR products and the primers 5'GGAATTCCATGAATGCATCCCTGGAGAG3' (SEQ ID NO: 10) and 5'CCCTCGAGTCATAAGGCCACTCGGATGC3' (SEQ ID NO: 11). The PCR conditions were 94°C for 15 s, 45°C for 20 s, 72°C for 2 min., for 25 cycles. The 1.6 kb secondary PCR product was treated with *EcoRI* and *XhoI* and subcloned into pGEX4T-3 (Pharmacia). The protein is expressed in *E. coli* strain LE392 at moderate levels upon IPTG induction and is soluble.

Independent confirmation of the existence of a mRNA corresponding to the full-length cDNA produced was carried out by RT-PCR. cDNA was prepared from 100 ng of human skeletal muscle polyA RNA (Clontech cat. no. 6541-1) using random hexamer primers and MoMLV reverse transcriptase. One twentieth of the cDNA was used as template in a PCR reaction containing the following primers sets: A1/P1, A2/P1, A2/P2, and A2/7-2 (A1: 5'GTGAGCTGACCCTGCTGGAG3' (SEQ ID NO: 12); A2: 5'AGACCTAAGCCTGTTTGCTCC3' (SEQ ID NO: 13); P1: 5'ACCGTGTCTGACTGCATGCT3' (SEQ ID NO: 14); P2: 5'TGAAGTTCCTTGGTGGAGC3' (SEQ ID NO: 15); 7-2: 5'CCCGTGAAACAGTGCTGTG3' (SEQ ID NO: 16)). The expected 288 bp, 203 bp, 954 bp and 1461 bp PCR fragments were

observed, respectively. The PCR conditions were 94°C for 15 s, 70°C to 50°C in 0.5°C increment/cycle for 20 s, 72°C for 2 min., for 40 cycles. Control reactions containing either no template or the 1.6 kb recombined cDNA produced above gave either no PCR product or the expected fragments.

Sequence analysis of the full-length cDNA revealed a 1608 nucleotide open reading frame (SEQ ID NO: 1) encoding a 536 amino acid protein (SEQ ID NO: 2) with a predicted molecular mass of 59 kDa, starting with an ATG at position 289 and terminating with a TGA at position 1897 of SEQ ID NO: 1.

GenBank searches using the BLASTX and BLASTP algorithms with the full-length cDNA sequence or with the deduced amino acid sequence were carried out to identify homologous entries. The search results indicated that the isolated full-length cDNA is an alternatively spliced isoform of Grb-IR (Liu et al., *supra*, GenBank Accession U34355 (SEQ ID NO: 17 and SEQ ID NO: 18)) designated as GrbIR-1, and is a member of the Grb10/Grb7 family of SH2 adapter proteins. See Fig. 1 for a multiple sequence alignment of GrbIR-1, Grb-IR, murine Grb10 and human Grb7.

An alignment of Grb-IR and GrbIR-1 using the GAP algorithm is shown in Fig. 2 (top, GrbIR-1; bottom, Grb-IR). The overall amino acid identity was 99.6% with one gap. GrbIR-1 contains an insert which restores an incomplete pleckstrin homology (PH) domain in Grb-IR and GrbIR-1 contains a shortened N-terminus when compared with Grb-IR. The regions other than the C-terminal SH2 domain and the PH domain did not show significant homologies to other database entries.

Example 2Tissue Distribution of GrbIR-1

Northern blots of tissue mRNA were conducted to determine the tissue distribution of *grbIR-1* gene transcription. The cDNA insert was amplified by PCR using the primers T3 and T7 and the 3.5 kb product was purified. Twenty-five ng of the PCR product was radiolabelled with [32P]-dATP using random hexamer primers and used to probe human multiple tissue Northern blots (Clontech cat. nos. 7760-1 and 7759-1). The membranes were washed at high stringency and exposed for 6 hrs to a storage phosphor screen (Molecular Dynamics) for visualization. Expression of the corresponding mRNA was largely ubiquitous and variable in level in heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testes, ovaries, small intestine and colon, although absent from peripheral blood leukocytes. The mRNA is approximately 5.6 kb in length. Highest expression was observed in heart, brain, skeletal muscle, and pancreas. Two additional transcripts are observed in skeletal muscle, of 4.8 and 3.1 kb. These may correspond to additional protein isoforms in this tissue.

The present invention may be embodied in other specific forms without departing from the spirit or essential attributes thereof, and, accordingly, reference should be made to the appended claims, rather than to the foregoing specification, as indicating the scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: SmithKline Beecham Corporation and Harvard University
- (ii) TITLE OF THE INVENTION: GROWTH FACTOR RECEPTOR-BINDING INSULIN RECEPTOR
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: SmithKline Beecham Corporation
 - (B) STREET: 709 Swedeland Road
 - (C) CITY: King of Prussia
 - (D) STATE: PA
 - (E) COUNTRY: USA
 - (F) ZIP: 19406
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ Version 1.5
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE: 09-JULY-1996
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Baumeister, Kirk
- (B) REGISTRATION NUMBER: 33.833
- (C) REFERENCE/DOCKET NUMBER: P50508P

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 610-270-5096
- (B) TELEFAX:
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2505 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGGCGCAACT TTGGCTCCCA GGGAACAAAC ATCCTCCTTC TAAGTGGTAG ATGTGGGTGA	60
GCTGACCCTG CTGGAGTCTG TCCCCTGGGC TACCCTCTGC TTCCCCCAT TGTGAGTGGT	120
CCGTGAAGCA CAGCGTTGAC CAGACCTAAG CCTGTTTGCT CCCAGGACAA GGTGGAGCAG	180
ACACCTCGCA GTCAACAAGA CCCGGCAGGA CCAGGACTCC CCGCACAGTC TGACCGACTT	240
GCGAATCACC AGGAGGATGA TGTGGACCTG GAAGCCCTGG TGAACGATAT GAATGCATCC	300
CTGGAGAGCC TGTACTCGGC CTGCAGCATG CAGTCAGACA CGGTGCCCCCT CCTGCAGAAT	360
GGCCAGCATG CCCGCAGCCA GCCTCGGGCT TCAGGCCCTC CTCGGTCCAT CCAGCCACAG	420
GTGTCCCCGA GGCAGAGGGT GCAGCGCTCC CAGCCTGTGC ACATCCTCGC TGTCAGGCGC	480
CTTCAGGAGG AAGACCAGCA GTTTAGAACC TCATCTCTGC CGGCCATCCC CAATCCTTTT	540
CCTGAACTCT GTGGCCCTGG GAGCCCCCCT GTGCTCACGC CGGGTTCTTT ACCTCCGAGC	600
CAGGCCGCCG CAAAGCAGGA TGTAAAGTC TTTAGTGAAG ATGGGACAAG CAAAGTGGTG	660
GAGATTCTAG CAGACATGAC AGCCAGAGAC CTGTGCCAAT TGCTGGTTTA CAAAAGTCAC	720
TGTGTGGATG ACAACAGCTG GACACTAGTG GAGCACCACC CGCACCTAGG ATTAGAGAGG	780
TGCTTGGAAG ACCATGAGCT GGTGGTCCAG GTGGAGAGTA CCATGGCCAG TGAGAGTAA	840
TTTCTATTCA GGAAGAATTA CGCAAATAC GAGTTCTTTA AAAATCCCAT GAATTCCTTC	900

CCAGAACAGA	TGTTACTTG	GTGCCAGCAG	TCAAATGGCA	GTCAAACCCA	GCTTTTGCAG	960
AATTTTCTGA	ACTCCAGTAG	TTGTCTTGAA	ATTCAAGGGT	TTTTGCATGT	GAAAGAGCTG	1020
GGAAAGAAAT	CATGGAAAAA	GCTGTATGTG	TGTTTGC GGA	GATCTGGCCT	TTATTGCTCC	1080
ACCAAGGGAA	CTTCAAAGGA	ACCCAGACAC	CTGCAGCTGC	TGGCCGACCT	GGAGGACAGC	1140
AACATCTTCT	CCCTGATCGC	TGGCAGGAAG	CAGTACAACG	CCCCTACAGA	CCACGGGCTC	1200
TGCATAAAGC	CAAACAAAGT	CAGGAATGAA	ACTAAAGAGC	TGAGGTTGCT	CTGTGCAGAG	1260
GACGAGCAAA	CCAGGACGTG	CTGGATGACA	GCGTTCAGAC	TCCTCAAGTA	TGAAATGCTC	1320
CTTTACCAGA	ATTACCGAAT	CCCTCAGCAG	AGGAAGGCCT	TGCTGTCCCC	GTTCTCGACG	1380
CCAGTGC GCA	GTGTCTCCGA	GAAC TCCCTC	GTGGCAATGG	ATTTTCTGG	GCAAACAGGA	1440
CGCGTGATAG	AGAATCCGGC	GGAGGCC CAG	AGCGCAGCCC	TGGAGGAGGG	CCACGCCTGG	1500
AGGAAGCGAA	GCACACGGAT	GAACATCCTA	GGTAGCCAAA	GTCCCTCCA	CCCTTCTACC	1560
CTAAGTACAG	TGATTACAG	GACACAGCAC	TGGTTTCAGC	GGAGGTTCTC	CAGGGAGGAA	1620
TCCCACAGGA	TCATTAAACA	GCAAGGGCTC	GTGGATGGGC	TTTTTCTCCT	CCGTGACAGC	1680
CAGAGTAATC	CAAAGGCATT	TGTACTCACA	CTGTGTCATC	ACCAGAAAAT	TAAAAATTTT	1740
CAGATCTTAC	CTTGCGAGGA	CGACGGGCAG	ACGTTCTTCA	GCCTAGATGA	CGGGAACACC	1800
AAATTCTCTG	ACCTGATCCA	GCTGGTTGAC	TTTTACCAGC	TGAACAAAGG	AGTCCTGCCT	1860
TGCAAAC TCA	AGCACCACTG	CATCCGAGTG	GCCTTATGAC	CGCAGATGTC	CTCTCGGCTG	1920
AAGACTGGAG	GAAGTGAACA	CTGGAGTGAA	GAAGCGGTCT	GTGCGTTGGT	GAAGAACACA	1980
CATCGATTCT	GCACCTGGGG	ACCCAGAGCG	AGATGGGTTT	GTTCCGGTGCC	AGCCTACCAA	2040
GATTGACTAG	TTTGTTGGAC	TTAAACGACG	ATTGCTGCT	GTGAACCCAG	CAGGGTCGCC	2100
TCCCTCTGCG	TCGGNCAAAT	TGGGGAGGGC	ATGGAAGATC	CAGCGGAAAG	TTGAAAATAA	2160
ACTGGAATGA	TCATCTTGGC	TTGGGCCGCT	TAGGAACAAG	AACCGGAGAG	AAGTGATTGG	2220
AAATGAACTC	TTGCCCTGGA	ATAATCTTGA	CAATTAAAAC	TGATATGTTT	ACTTTTTTTG	2280
TATTGATCAC	TTTTTTGGAC	TCCTTCTTTG	TTTCAATAT	TGTATTCAGC	CTATTGTAGG	2340
AGGGGGATGT	GGCGTTTCAA	CTCATATAAT	ACAGAAAGAG	TTTTGGAATG	GGCAGATTTT	2400
AAACTGAATA	TGGGTCCCCA	AATGTTCCCA	GAGGGTCCCT	CACAACCTCT	GNCGACTACC	2460
ACGGTG TNGG	ATT CAGCTCC	CAAATGACAA	ACCCAGNCCT	TCCCA		2505

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 536 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Asn Ala Ser Leu Glu Ser Leu Tyr Ser Ala Cys Ser Met Gln Ser
 1           5           10           15
Asp Thr Val Pro Leu Leu Gln Asn Gly Gln His Ala Arg Ser Gln Pro
      20           25           30
Arg Ala Ser Gly Pro Pro Arg Ser Ile Gln Pro Gln Val Ser Pro Arg
      35           40           45
Gln Arg Val Gln Arg Ser Gln Pro Val His Ile Leu Ala Val Arg Arg
      50           55           60
Leu Gln Glu Glu Asp Gln Gln Phe Arg Thr Ser Ser Leu Pro Ala Ile
65           70           75           80
Pro Asn Pro Phe Pro Glu Leu Cys Gly Pro Gly Ser Pro Pro Val Leu
      85           90           95
Thr Pro Gly Ser Leu Pro Pro Ser Gln Ala Ala Ala Lys Gln Asp Val
      100          105          110
Lys Val Phe Ser Glu Asp Gly Thr Ser Lys Val Val Glu Ile Leu Ala
      115          120          125
Asp Met Thr Ala Arg Asp Leu Cys Gln Leu Leu Val Tyr Lys Ser His
      130          135          140
Cys Val Asp Asp Asn Ser Trp Thr Leu Val Glu His His Pro His Leu
      145          150          155          160
Gly Leu Glu Arg Cys Leu Glu Asp His Glu Leu Val Val Gln Val Glu
      165          170          175
Ser Thr Met Ala Ser Glu Ser Lys Phe Leu Phe Arg Lys Asn Tyr Ala
      180          185          190
Lys Tyr Glu Phe Phe Lys Asn Pro Met Asn Phe Phe Pro Glu Gln Met
      195          200          205
Val Thr Trp Cys Gln Gln Ser Asn Gly Ser Gln Thr Gln Leu Leu Gln
      210          215          220
Asn Phe Leu Asn Ser Ser Ser Cys Pro Glu Ile Gln Gly Phe Leu His
      225          230          235          240
Val Lys Glu Leu Gly Lys Lys Ser Trp Lys Lys Leu Tyr Val Cys Leu
      245          250          255
Arg Arg Ser Gly Leu Tyr Cys Ser Thr Lys Gly Thr Ser Lys Glu Pro
      260          265          270
Arg His Leu Gln Leu Leu Ala Asp Leu Glu Asp Ser Asn Ile Phe Ser
      275          280          285
Leu Ile Ala Gly Arg Lys Gln Tyr Asn Ala Pro Thr Asp His Gly Leu
      290          295          300
Cys Ile Lys Pro Asn Lys Val Arg Asn Glu Thr Lys Glu Leu Arg Leu
      305          310          315          320

```



```

Leu Cys Ala Glu Asp Glu Gln Thr Arg Thr Cys Trp Met Thr Ala Phe
      325                      330                      335
Arg Leu Leu Lys Tyr Glu Met Leu Leu Tyr Gln Asn Tyr Arg Ile Pro
      340                      345                      350
Gln Gln Arg Lys Ala Leu Leu Ser Pro Phe Ser Thr Pro Val Arg Ser
      355                      360                      365
Val Ser Glu Asn Ser Leu Val Ala Met Asp Phe Ser Gly Gln Thr Gly
      370                      375                      380
Arg Val Ile Glu Asn Pro Ala Glu Ala Gln Ser Ala Ala Leu Glu Glu
      385                      390                      395                      400
Gly His Ala Trp Arg Lys Arg Ser Thr Arg Met Asn Ile Leu Gly Ser
      405                      410                      415
Gln Ser Pro Leu His Pro Ser Thr Leu Ser Thr Val Ile His Arg Thr
      420                      425                      430
Gln His Trp Phe His Gly Arg Phe Ser Arg Glu Glu Ser His Arg Ile
      435                      440                      445
Ile Lys Gln Gln Gly Leu Val Asp Gly Leu Phe Leu Leu Arg Asp Ser
      450                      455                      460
Gln Ser Asn Pro Lys Ala Phe Val Leu Thr Leu Cys His His Gln Lys
      465                      470                      475                      480
Ile Lys Asn Phe Gln Ile Leu Pro Cys Glu Asp Asp Gly Gln Thr Phe
      485                      490                      495
Phe Ser Leu Asp Asp Gly Asn Thr Lys Phe Ser Asp Leu Ile Gln Leu
      500                      505                      510
Val Asp Phe Tyr Gln Leu Asn Lys Gly Val Leu Pro Cys Lys Leu Lys
      515                      520                      525
His His Cys Ile Arg Val Ala Leu
      530                      535

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2420 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATAATTCTC	AAATTTTCT	TACTTACCTA	AATATAAGCT	AATTTCTATA	ACTCTAATTC	60
CTCAAAAGGT	ACTCCCTCTC	TCTCTCTCTC	TCTCTCCCTC	TCTCCTAGCA	CCTGCTGCTC	120
AGTAGGAAGG	GCAAGAGCAA	TTGAGGCCG	GTGCATTGTG	AGGAGTCTCC	ACCCCTCCTC	180
CTGCGCTTCC	TTCTCCAGGG	AGCCTCTCAG	GCCGCCCTCA	CCTGCCCGAG	ATAATTTTAG	240
TTTCCCTGGG	CCTGGAATCT	GGATACGCAG	GGCCTCGCTC	TATATTCTCC	CGCCTCAACA	300
TTCCAAAGGC	GGGATAGCCT	TTCTACCATC	TGTAGAGAAG	AGAGAAAGGA	TTCGAAATCA	360
AATCCAAGTG	TCTGGGATCT	CTAGACAGAG	CCAGACTTTG	GGCCGGGTGT	CCGGCTCCTT	420
CTGTTGGAGG	TGCTCCAGGT	GCCATGGAAC	TGGATCTGAG	CCCRACTCAT	CTCAGCAGCT	480
CCCCAGAAGA	TGTGTGCCCA	ACTCCTGCTA	CCCCTCCTGA	GACTCCTCCG	CCCCCTGATA	540
ACCCCTCGCC	AGGGGATGTG	AAGCGGTGCG	AGCCTTTGCC	CATCCCCAGC	AGCAGGAAAC	600
TTGAGAAGA	GGAGTTTCAG	GCAACCTCTC	TGCCCTCAT	CCCCAACCCC	TTCCCTGAGC	660
TCTGCAGCCC	ACCTTCACAG	AAACCCATTC	TTGGTGGTTC	CTCCGGTGCA	AGGGGGTTGC	720
TTCTCGAGA	CTCCAGCCGC	CTCTGTGTGG	TGAAGTGTA	CAGTGAGGAT	GGGGCTGCC	780
GGTCTGTGGA	GGTGGCAGCG	GGGCCACAG	CTCGTCACGT	GTGTGAGATG	CTGGTACAAC	840
GAGCTCACGC	CCTGAGCGAC	GAGAGCTGGG	GACTAGTGGA	ATCCCCCCCC	TACCTGGCAC	900
TGGAGCGGGG	TCTGGAGGAC	CATGAATTTG	TGGTGAAGT	GCAGGAGGCC	TGGCCTGTGG	960
GTGGAGATAG	CCGCTTCATC	TTCCGTAAAA	ACTTCGCCAA	GTATGAACTA	TTCAAGAGCC	1020
CCCCACACAC	CCTGTTTCCA	GAAAAGATGG	TCTCGAGCTG	TCTGGATGCA	CAAACAGGCA	1080
TATCCCATGA	AGACCTCATC	CAGAACTTCC	TGAACGCTGG	CAGCTTCCCT	GAGATCCAGG	1140
GCTTCTGCA	GCTGCGGGGA	TCAGGCCGGG	GGTCAGGTCG	AAAGCTTTGG	AAACGTTTCT	1200
TCTGCTTTCT	CGGTCGATCT	GGCCTCTACT	ACTCTACCAA	GGGTACCTCC	AAGGACCCCA	1260
GACACCTACA	GTATGTGGCA	GATGTGAATG	AGTCCAATGT	CTATGTGGTG	ACCCAGGGCC	1320
GCAAGCTGTA	TGGGATGCCC	ACTGACTTCG	GCTTCTGTGT	CAAGCCCAAC	AAGCTTCGAA	1380
ACGGCCACAA	GGGGCTCCAC	ATCTTCTGCA	GTGAGGATGA	GCAGAGTCGG	ACCTGCTGGC	1440
TGGCTGCCTT	CCGGCTCTTC	AAGTACGGGG	TACAGCTATA	TAAGAATTAT	CAGCAGGCCC	1500
AGTCTCGTCA	CCTGCGCCTA	TCCTATTTGG	GGTCTCCACC	CTTGAGGAGC	GTCTCAGACA	1560
ATACCCTAGT	GGCTATGGAC	TTCTCTGGCC	ATGCGGGGCG	TGTCATTGAT	AACCCCGGG	1620
AAGCTCTGAG	TGCCGCCATG	GAGGAGGCC	AGGCCTGGAG	GAAGAAGACA	AACCACCGTC	1680
TGAGCCTGCC	CACCACATGC	TCTGGCTCGA	GCCTCAGCGC	AGCCATTTCAT	CGCACCACAGC	1740
CCTGGTTTCA	TGGACGCATC	TCTCGGGAGG	AGAGCCAGCG	GCTAATTGGA	CAGCAGGGCC	1800
TGGTGGATGG	TGTGTTCCCTG	GTCCGGGAGA	GCCAGAGGAA	CCCACAGGGC	TTTGTCTGT	1860
CCTTGTGCCA	TCTGCAGAAA	GTCAAGCATT	ATCTCATTTT	GCCAAGTGAA	GATGAAGGTT	1920
GCCTTTACTT	CAGCATGGAT	GAGGGCCAGA	CCCGTTTCAC	AGACCTGCTG	CAGCTGGTAG	1980
AATTCACCA	GCTGAACCGA	GGCATCCTGC	CCTGCCTGCT	GCGCCACTGC	TGTGCCCCGTG	2040
TGGCCCTCTG	AGGCCGCACA	AGCTACTGCA	GCCATGGGTT	TGCCTACCAC	CCTTCTGTCC	2100
TGTGGACTCG	GTGCAGGTGG	GTGGGTGGT	AAACAGTGGA	AGAGCTCCCC	CCCCCAATTT	2160
TATCCCATTT	TTTTTAACCT	CTCTCAACCA	GTGAAACATC	CCCTAACCCCT	GTCCATCCCT	2220
GACTCCTGTC	CCCAAGGGAG	GCATTGTGGT	CCTGTCCCCT	TGGTAGAGCT	CCTGAGGTAC	2280
TGTTCCAGTG	AGGGGCATTA	TGAGAGGAGC	GGGGCAGCCC	AGGAGGTCTC	ATACCCACCC	2340

CATAATCTGT ACAGACTGAG AGGCCAGTTG ATCTGCTCTG TTTTATACCA GTAACAATAA 2400
AGATTATTTT TTGATACAAA 2420

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GATTGGTGGC GACGACTCC

19

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCCGTGAAAC CAGTGCTGTG

20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTAATACGAC TCACTATAGG GC

22

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGTAGCCAAA GTCCCCTCCA

20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATTGGTGGC GACGACTCC

19

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGAGGGGAC TTTGGCTACC

20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGAATTCCAT GAATGCATCC CTGGAGAG

28

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCCTCGAGTC ATAAGGCCAC TCGGATGC

28

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTGAGCTGAC CCTGCTGGAG

20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGACCTAAGC CTGTTTGCTC C

21

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACCGTGTCTG ACTGCATGCT

20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TGAAGTTCCC TTGGTGGAGC

20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 0 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCCGTGAAACCAAGTGCTGTG

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2070 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AAATGTAATT	TGAAGAAGGC	AGAAGGAACC	CATGGCTTTA	GCCGGCTGCC	CAGATTCCTT	60
TTTGCACCAT	CCGTACTACC	AGGACAAGGT	GGAGCAGACA	CCTCGCAGTC	AACAAGACCC	120
GGCAGGACCA	GGACTCCCCG	CACAGTCTGA	CCGACTTGCG	AATCACCAGG	AGGATGATGT	180
GGACCTGGAA	GCCCTGGTGA	ACGATATGAA	TGCATCCCTG	GAGAGCCTGT	ACTCGGCCTG	240
CAGCATGCAG	TCAGACACGG	TGCCCCCTCT	GCAGAAATGGC	CAGCATGCCC	GCAGCCAGCC	300
TCGGGGCTTCA	GGCCCTCCTC	GGTCCATCCA	GCCACAGGTG	TCCCCGAGGC	AGAGGGTGCA	360
GCGCTCCAG	CCTGTGCACA	TCCTCGCTGT	CAGGCGCCTT	CAGGAGGAAG	ACCAGCAGTT	420
TAGAACCTCA	TCTCTGCCGG	CCATCCCCAA	TCCTTTTCCT	GAACTCTGTG	GCCCTGGGAG	480
CCCCCCTGTG	CTCACGCCGG	GTTCTTTACC	TCCGAGCCAG	GCCGCCGCAA	AGCAGGATGT	540
TAAAGTCTTT	AGTGAAGATG	GGACAAGCAA	AGTGGTGGAG	ATTCTAGCAG	ACATGACAGC	600
CAGAGACCTG	TGCCAATTGC	TGGTTTACAA	AAGTCAC'TGT	GTGGATGACA	ACAGCTGGAC	660
ACTAGTGGAG	CACCACCCGC	ACCTAGGATT	AGAGAGGTGC	TTGGAAGACC	ATGAGCTGGT	720
GGTCCAGGTG	GAGAGTACCA	TGGCCAGTGA	GAGTAAATTT	CTATTCAGGA	AGAATTACGC	780
AAAATACGAG	TTCTTTAAAA	ATCCCATGAA	TTTCTTCCCA	GAACAGATGG	TTACTTGGTG	840
CCAGCAGTCA	AATGGCAGTC	AAACCCAGCT	TTTGCAGGAA	CCCAGACACC	TGCAGCTGCT	900
GGCCGACCTG	GAGGACAGCA	ACATCTTCTC	CCTGATCGCT	GGCAGGAAGC	AGTACAACGC	960
CCCTACAGAC	CACGGGCTCT	GCATAAAGCC	AAACAAAGTC	AGGAATGAAA	CTAAAGAGCT	1020
GAGGTTGCTC	TGTGCAGAGG	ACGAGCAAAC	CAGGACGTGC	TGGATGACAG	CGTTCAGACT	1080
CCTCAAGTAT	GGAATGCTCC	TTTACCAGAA	TTACCGAATC	CCTCAGCAGA	GGAAGGCCTT	1140
GCTGTCCCCG	TTCTCGACGC	CAGTGCAGCAG	TGTCTCCGAG	AACTCCCTCG	TGGCAATGGA	1200
TTTTTCTGGG	CAAACAGGAC	GCGTGATAGA	GAATCCGGCG	GAGGCCCAGA	GCGCAGCCCT	1260
GGAGGAGGGC	CACGCCTGGA	GGAAGCGAAG	CACACGGATG	AACATCCTAG	GTAGCCAAAG	1320
TCCCTCCAC	CCTTCTACCC	TAAGTACAGT	GATTACACAGG	ACACAGCACT	GGTTTCACGG	1380
GAGGATCTCC	AGGGAGGAAT	CCCACAGGAT	CATTAAACAG	CAAGGGCTCG	TGGATGGGCT	1440
TTTTCTCCTC	CGTGACAGCC	AGAGTAATCC	AAAGGCATTT	GTA'CTCACAC	TGTGTCATCA	1500
CCAGAAAATT	AAAAATTTCC	AGATCTTACC	TTGCGAGGAC	GACGGGCAGA	CGTTCTTCAG	1560
CCTAGATGAC	GGGAACACCA	AATTCTCTGA	CCTGATCCAG	CTGGTTGACT	TTTACCAGCT	1620
GAACAAAGGA	GTCTGCCTT	GCAAACTCAA	GCACCACTGC	ATCCGAGTGG	CCTTATGACC	1680
GCAGATGTCC	TCTCGGCTGA	AGACTGGAGG	AAGTGAACAC	TGGAGTGAAG	AAGCGGTCTG	1740
TGCGTTGGTG	AAGAACACAC	ATCGATTCTG	CACCTGGGGA	CCCAGAGCGA	GATGGGTTTG	1800
TTCCGGTGCCA	GCCGACCAAG	ATTGACTAGT	TTGTTGGACT	TAAACGACGA	TTTGCTGCTG	1860
TGAACCCAGC	AGGGTCGCCT	CCCTCTGCGT	CGGCCAAATT	GGGGAGGGCA	TGGAAGATCC	1920
AGCGGAAAGT	TGAAAATAAA	CTGGAATGAT	CATCTTGGCT	TGGGCCGCTT	AGGAACAAGA	1980
ACCGGAGAGA	AGTGATTGGA	AATGA'ACTCT	TGCCCTGGAA	TAATCTTGAC	AATTAAAACT	2040
GATATGTTTA	AAAAAAAAAA	AAAAAAA'ACT				2070

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 548 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

Met Ala Leu Ala Gly Cys Pro Asp Ser Phe Leu His His Pro Tyr Tyr
 1             5             10             15
Gln Asp Lys Val Glu Gln Thr Pro Arg Ser Gln Gln Asp Pro Ala Gly
      20             25             30
Pro Gly Leu Pro Ala Gln Ser Asp Arg Leu Ala Asn His Gln Glu Asp
      35             40             45
Asp Val Asp Leu Glu Ala Leu Val Asn Asp Met Asn Ala Ser Leu Glu
      50             55             60
Ser Leu Tyr Ser Ala Cys Ser Met Gln Ser Asp Thr Val Pro Leu Leu
      65             70             75             80
Gln Asn Gly Gln His Ala Arg Ser Gln Pro Arg Ala Ser Gly Pro Pro
      85             90             95
Arg Ser Ile Gln Pro Gln Val Ser Pro Arg Gln Arg Val Gln Arg Ser
      100            105            110
Gln Pro Val His Ile Leu Ala Val Arg Arg Leu Gln Glu Glu Asp Gln
      115            120            125
Gln Phe Arg Thr Ser Ser Leu Pro Ala Ile Pro Asn Pro Phe Pro Glu
      130            135            140
Leu Cys Gly Pro Gly Ser Pro Pro Val Leu Thr Pro Gly Ser Leu Pro
      145            150            155            160
Pro Ser Gln Ala Ala Ala Lys Gln Asp Val Lys Val Phe Ser Glu Asp
      165            170            175
Gly Thr Ser Lys Val Val Glu Ile Leu Ala Asp Met Thr Ala Arg Asp
      180            185            190
Leu Cys Gln Leu Leu Val Tyr Lys Ser His Cys Val Asp Asp Asn Ser
      195            200            205

```

Trp Thr Leu Val Glu His His Pro His Leu Gly Leu Glu Arg Cys Leu
 210 215 220
 Glu Asp His Glu Leu Val Val Gln Val Glu Ser Thr Met Ala Ser Glu
 225 230 235 240
 Ser Lys Phe Leu Phe Arg Lys Asn Tyr Ala Lys Tyr Glu Phe Phe Lys
 245 250 255
 Asn Pro Met Asn Phe Phe Pro Glu Gln Met Val Thr Trp Cys Gln Gln
 260 265 270
 Ser Asn Gly Ser Gln Thr Gln Leu Leu Gln Glu Pro Arg His Leu Gln
 275 280 285
 Leu Leu Ala Asp Leu Glu Asp Ser Asn Ile Phe Ser Leu Ile Ala Gly
 290 295 300
 Arg Lys Gln Tyr Asn Ala Pro Thr Asp His Gly Leu Cys Ile Lys Pro
 305 310 315 320
 Asn Lys Val Arg Asn Glu Thr Lys Glu Leu Arg Leu Leu Cys Ala Glu
 325 330 335
 Asp Glu Gln Thr Arg Thr Cys Trp Met Thr Ala Phe Arg Leu Leu Lys
 340 345 350
 Tyr Gly Met Leu Leu Tyr Gln Asn Tyr Arg Ile Pro Gln Gln Arg Lys
 355 360 365
 Ala Leu Leu Ser Pro Phe Ser Thr Pro Val Arg Ser Val Ser Glu Asn
 370 375 380
 Ser Leu Val Ala Met Asp Phe Ser Gly Gln Thr Gly Arg Val Ile Glu
 385 390 395 400
 Asn Pro Ala Glu Ala Gln Ser Ala Ala Leu Glu Glu Gly His Ala Trp
 405 410 415
 Arg Lys Arg Ser Thr Arg Met Asn Ile Leu Gly Ser Gln Ser Pro Leu
 420 425 430
 His Pro Ser Thr Leu Ser Thr Val Ile His Arg Thr Gln His Trp Phe
 435 440 445
 His Gly Arg Ile Ser Arg Glu Glu Ser His Arg Ile Ile Lys Gln Gln
 450 455 460
 Gly Leu Val Asp Gly Leu Phe Leu Leu Arg Asp Ser Gln Ser Asn Pro
 465 470 475 480
 Lys Ala Phe Val Leu Thr Leu Cys His His Gln Lys Ile Lys Asn Phe
 485 490 495
 Gln Ile Leu Pro Cys Glu Asp Asp Gly Gln Thr Phe Phe Ser Leu Asp
 500 505 510
 Asp Gly Asn Thr Lys Phe Ser Asp Leu Ile Gln Leu Val Asp Phe Tyr
 515 520 525

WO 98/01475

PCT/US96/11452

Gln Leu Asn Lys Gly Val Leu Pro Cys Lys Leu Lys His His Cys Ile

530

535

540

Arg Val Ala Leu

545

CLAIMS

1. An isolated polynucleotide selected from the group consisting of:
 - (a) a polynucleotide encoding human GrbIR-1 having
5 the nucleotide sequence as set forth in SEQ ID NO:1 from nucleotide 289 to 1897;
 - (b) a polynucleotide capable of hybridizing to the complement of a polynucleotide according to (a) under moderately stringent hybridization conditions and which
10 encodes a functional human GrbIR-1; and
 - (c) a degenerate polynucleotide according to (a) or (b).
2. An isolated polynucleotide having the nucleotide sequence as set forth in SEQ ID NO:1.
- 15 3. A functional polypeptide encoded by the polynucleotide of claim 1.
4. The functional polypeptide of claim 3 which is human GrbIR-1 having the amino acid sequence set forth in SEQ ID NO:2.
- 20 5. The polynucleotide of claim 1 which is DNA.
6. The polynucleotide of claim 5 which is genomic DNA.
7. The polynucleotide of claim 1 which is RNA.
8. A vector comprising the DNA of claim 5.
- 25 9. A recombinant host cell comprising the vector of claim 8.
10. A method for preparing essentially pure human GrbIR-1 protein comprising culturing the recombinant host cell of claim 9 under conditions promoting
30 expression of the protein and recovering the expressed protein.
11. Human GrbIR-1 produced by the process of claim 10.
12. An antisense oligonucleotide comprising a
35 sequence which is capable of binding to the polynucleotide of claim 1.
13. A modulator of the polypeptide of claim 3.

14. The modulator of claim 13 which is a peptide.
15. The modulator of claim 13 which is a small organic molecule.
16. The small organic molecule of claim 15 which is a peptidomimetic.
17. A method for assaying a medium for the presence of a substance that modulates GrbIR-1 activity by affecting the binding of GrbIR-1 to cellular binding partners comprising the steps of:
- 10 (a) providing a GrbIR-1 protein having the amino acid sequence of GrbIR-1 (SEQ ID NO:2) or a functional derivative thereof and a cellular binding partner or synthetic analog thereof;
- (b) incubating with a test substance which is suspected of modulating GrbIR-1 activity under conditions which permit the formation of a GrbIR-1 protein/cellular binding partner complex;
- 15 (c) assaying for the presence of the complex, free GrbIR-1 protein or free cellular binding partner; and
- 20 (d) comparing to a control to determine the effect of the substance.
18. GrbIR-1 protein modulating compounds identified by the method of claim 17.
- 25 19. A method for assaying for the presence of a substance that modulates GrbIR-1 activity by direct binding to GrbIR-1 protein comprising the steps of:
- (a) providing a labelled GrbIR-1 protein having the amino acid sequence of GrbIR-1 (SEQ ID NO:2) or a functional derivative thereof;
- 30 (b) providing solid-support-associated modulator candidates;
- (c) incubating a mixture of the labelled GrbIR-1 protein with the support-associated modulator candidates under conditions which can permit the formation of a GrbIR-1 protein/modulator candidate complex;
- 35

- (d) separating the solid support from free soluble labelled GrbIR-1 protein;
 - (e) assaying for the presence of solid support-associated labelled protein;
 - 5 (f) isolating the solid support complexed with labelled GrbIR-1 protein; and
 - (g) identifying the modulator candidate.
20. GrbIR-1 modulating compounds identified by the method of claim 19.
- 10 21. A method for the treatment of a patient having need to modulate GrbIR-1 activity comprising administering to the patient a therapeutically effective amount of the modulating compound of claims 18 or 20.
- 15 22. A pharmaceutical composition comprising the modulating compound of claims 18 or 20 and a pharmaceutically acceptable carrier.
23. A method of diagnosing conditions associated with GrbIR-1 protein deficiency which comprises:
- 20 (a) isolating a polynucleotide sample from an individual;
 - (b) assaying the polynucleotide sample and a polynucleotide encoding GrbIR-1 having the nucleotide sequence as set forth in SEQ ID NO:1 from nucleotide 289 to 1897; and
 - 25 (c) comparing differences between the polynucleotide sample and the GrbIR-1 polynucleotide, wherein any differences indicate mutations in the GrbIR-1 gene.
24. A method of treating conditions which are
- 30 related to insufficient GrbIR-1 protein function which comprises:
- (a) isolating cells from a patient deficient in GrbIR-1 protein function;
 - (b) altering the cells by transfecting the
 - 35 polynucleotide of claim 1 into the cells wherein a GrbIR-1 protein is expressed; and

(c) introducing the cells back to the patient to alleviate the condition.

25. A method of treating conditions which are related to insufficient GrbIR-1 protein function which
5 comprises administering the polynucleotide of claim 1 to a patient deficient in GrbIR-1 protein function wherein a GrbIR-1 protein is expressed and alleviates the condition.

26. A transgenic non-human animal capable of
10 expressing in any cell thereof the DNA of claim 5.


```

1
GrbIR-1 ..... 50
Grb-IR MALAGCPDSF LHHPPYQDKV EQTPRSQQDP AGPGLPAQSD RLANHQEDDV
mGrb10 .....
hGrb7 .....

51
GrbIR-1 ..... 100
Grb-IR ..... MN ASLESLSAC SMQS..DTVP LLQNGQHARS QPRASGPPRS
mGrb10 ..... MNNDIN SSVESLSAC NMQSDTDTAP LLEDGQHASN QGAASSR..
hGrb7 ..... MELDLSP HLSSSPEDL ..... W PAPGTPPGTP

101
GrbIR-1 IQPQVSPRQR VORSQPVHI. LAVRRLQEED QQFRTSSLPA IPNPFPELCG
Grb-IR IQPQVSPRQR VORSQPVHI. LAVRRLQEED QQFRTSSLPA IPNPFPELCG
mGrb10 GPPQASPRQK MORSQPVHI. L..RRLQEED QQLRTASLPA IPNPFPELTG
hGrb7 RPPDTLPPEE VKRSQPLLIP TTGRKLREEE R..RATSLPS IPNPFPELCS

151
GrbIR-1 ..PGSPPVLT PGS..PPSQ AAAKQ ..... 200
Grb-IR ..PGSPPVLT PGS..PPSQ AAAKQ .....
mGrb10 AAPGSPPSVA PSSLPPPSQ PPAKHCGRCE KWIPGENTRG NGKRKIWRWQ
hGrb7 PPSQSPILGG PSSARGLLPR DASRPHV...

201
GrbIR-1 ..... 250
Grb-IR .....
mGrb10 FPPGFQLSKL TRPGLWTKT ARFSKKQPKN QCPTDTVNPV ARMPTSQMEK
hGrb7 .....

251
GrbIR-1 ..... 300
Grb-IR ..... DVKVF SEDGTSKVVE ILADMTARDL CQLLVYKSHC VDDNSWTLVE
mGrb10 LRLRKDVKVF SEDGTSKVVE ILADMTARDL CQLLVYKSHC VDDNSWTLVE
hGrb7 ..... VKVY SEDGACRSVE VAAGATARHV CEMLVQRAHA LSDETGWGLVE

301
GrbIR-1 HHPHLGLERC LEDHELVVQV EST..MASES KFLFRKNYAK YEFFK.NPMN 350
Grb-IR HHPHLGLERC LEDHELVVQV EST..MASES KFLFRKNYAK YEFFK.NPMN
mGrb10 HHPQLGLERC LEDHEIVVQV EST..MPSES KFLFRKNYAK YEFFK.NPVN
hGrb7 CHPHLALERG LEDHESVVEV QAAWPVGGDS RFVFRKNFAK YELFKSSPHS

351
GrbIR-1 FFPEQMVTWC QQSNG..SQT QLLQNFLNSS SCPEIQGFLH VKELGKKSWK 400
Grb-IR FFPEQMVTWC QQSNG..SQT QLLQ.....
mGrb10 FFPDQMVNWC QQSNG..GQA QLLQNFLNTS SCPEIQGFLQ VKEVGRKSWK
hGrb7 LFPEKMVSSC LDAHTGISHE DLIQNFLNAG SFPEIQGFLQ LRGSGRKLWK

401
GrbIR-1 KLYVCLRRSG LYCSTKGTSK EPRHLQLLAD LEDSNIFSLI AGRKQYNAPT 450
Grb-IR ..... EPRHLQLLAD LEDSNIFSLI AGRKQYNAPT
mGrb10 KLYVCLRRSG LYYSTKGTSK EPRHLQLLAD LEESIFLYI AGKKQYNAPN
hGrb7 RFFCFLRRSG LYYSTKGTSK DPRHLQYVAD VNESNVYVVT QGRKLYGMPT

451
GrbIR-1 DHGLCIKPNK VRNETKELRL LCAEDEQTRT CWMTAFRLK YEMLLYQNYR 500
Grb-IR DHGLCIKPNK VRNETKELRL LCAEDEQTRT CWMTAFRLK YGMLLYQNYR
mGrb10 EHGMCIKPNK AKTEMKELRL LCAEDEQIRT CWMTAFRLK YGMLLYQNYR
hGrb7 DFGFCVKPNK LRNGHKGLRI FCSEDEQSRT CWLAFAFLFK YGVQLYKNYQ

```

FIGURE 1A/2

	501		550
GrbIR-1	IPQQRKALLS	PF.STPVRSV	SENSLVAMDF SGQTGRVIEN PAEAQSAALE
Grb-IR	IPQQRKALLS	PF.STPVRSV	SENSLVAMDF SGQTGRVIEN PAEAQSAALE
mGrb10	IP.QRKGLPP	PF.NAPMRSV	SENSLVAMDF SGQIGRVIDN PAEAQSAALE
hGrb7	QAQSRHLHPS	CLGSPPLRSA	SDNTLVAMDF SGHAGRVIEH PREALSVALE
	551		600
GrbIR-1	EGHAWRKRST	RMNILGSQSP	LHPSTLSTVI HRTQHWFHGR FSREESHRII
Grb-IR	EGHAWRKRST	RMNILGSQSP	LHPSTLSTVI HRTQHWFHGR ISREESHRII
mGrb10	EGHAWRNGST	RMNILSSQSP	LHPSTLNAVI HRTQHWFHGR ISREESHRII
hGrb7	EAQAWRKKTN	HR..LSLPMP	ASGTSLSAAI HRTQLWFHGR ISREESQRLI
	601		650
GrbIR-1	KQQGLVDGLF	LLRDSQSNPK	AFVLTLCCHQ KIKNFQILPC EDDGQTFEFL
Grb-IR	KQQGLVDGLF	LLRDSQSNPK	AFVLTLCCHQ KIKNFQILPC EDDGQTFEFL
mGrb10	KQQGLVDGLF	LLRDSQSNPK	AFVLTLCCHQ KIKNFQILPC EDDGQTFEFL
hGrb7	GQQGLVDGLF	LVRESQRNPQ	GFVLSLCHLQ KVKHYLILPS EEGRLYFSM
	651		687
GrbIR-1	DDGNTKFSDL	IQLVDFYQLN	KGVLPCCLKH HCIRVAL
Grb-IR	DDGNTKFSDL	IQLVDFYQLN	KGVLPCCLKH HCIRVAL
mGrb10	DDGNTKFSDL	IQLVDFYQLN	KGVLPCCLKH HCIRVAL
hGrb7	DDGQTRFTDL	LQLVEFHQLN	RGILPCLLRH CCTRVAL

FIGURE 1B/2

1MNASLESLSACSMQSDTVPLLQNGQHARSQPRASGPPRSIQ 42
|||||
51 DLEALVNDMNASLESLSACSMQSDTVPLLQNGQHARSQPRASGPPRSIQ 100
|||||
43 PQVSPRQVRORSQPVHILAVRRLQEEDQQFRTSSLPAIPNPFPELCGPGS 92
|||||
101 PQVSPRQVRORSQPVHILAVRRLQEEDQQFRTSSLPAIPNPFPELCGPGS 150
|||||
93 PPVLTPGSLPPSQAAAKQDVKFSESGTSGVVEILADMTARDLCQLLVYK 142
|||||
151 PPVLTPGSLPPSQAAAKQDVKFSESGTSGVVEILADMTARDLCQLLVYK 200
|||||
143 SHCVDNDSWTLVEHHPHLGLERCLEDHELVVQVESTMASESKEFLFRKNYA 192
|||||
201 SHCVDNDSWTLVEHHPHLGLERCLEDHELVVQVESTMASESKEFLFRKNYA 250
|||||
193 KYEFFKNPMNFFPEQMVTWCQQSNGSQTOLLQNFLNSSSCPETIQGFLHVK 242
|||||
251 KYEFFKNPMNFFPEQMVTWCQQSNGSQTOLLQ..... 282
|||||
243 ELGKKSWKLYVCLRRSGLYCSTKGTSKEPRHLQLLADLEDSNIFSLIAG 292
|||||
283EPRHLQLLADLEDSNIFSLIAG 304
|||||
293 RKQYNAPT DHGLCIKPNKVRNETKELRLLCAEDEQTRTCWMTAFRLKYE 342
|||||
305 RKQYNAPT DHGLCIKPNKVRNETKELRLLCAEDEQTRTCWMTAFRLKYG 354
|||||
343 MLLYQNYRIPQQRKALLSPFSTPVRSVSENSLVAMDFSGQTGRVIENPAE 392
|||||
355 MLLYQNYRIPQQRKALLSPFSTPVRSVSENSLVAMDFSGQTGRVIENPAE 404
|||||
393 AQSAALEEGHAWRKSTRMNILGSQSPLHPSTLSTVIHRTQHWFHGRFSR 442
|||||
405 AQSAALEEGHAWRKSTRMNILGSQSPLHPSTLSTVIHRTQHWFHGRISR 454
|||||
443 EESHRIIKQOGLVDGLFLLRDSQSNPKAFVLTLCCHQKIKNFQILPCEDD 492
|||||
455 EESHRIIKQOGLVDGLFLLRDSQSNPKAFVLTLCCHQKIKNFQILPCEDD 504
|||||
493 GQTFSLDDGNTKFSDLIQLVDFYQLNKGVLPCCLKKHCIRVAL 536
|||||
505 GQTFSLDDGNTKFSDLIQLVDFYQLNKGVLPCCLKKHCIRVAL 548
|||||

FIGURE 2/2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/11452

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : CO7K 14/71, 14/62, 14/475; C12P 21/02; C12N 1/21, 5/10, 15/09
US CL : 435/69.1, 240.1, 243, 320.1; 530/350, 303; 536/23.5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 240.1, 243, 320.1; 530/350, 303; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- A	US 5,434,064 A (SCHLESSINGER ET AL) 18 July 1995; see entire document, especially Figs. 19 and 36A-C; SEQ ID NO:7; column 14; and columns. 47-48.	1,3,5,7-9 ----- 2,4,6,10-11
X ---- A	LIU et al. Grb-IR: A SH2-domain-containing protein that binds to the insulin receptor and inhibits its function. Proc. Natl. Acad. Sci.. October 1995, Vol.92, pages 10287-10291, especially Fig. 1 and "Materials and Methods" section.	1,3,5,7-9 ----- 2,4,6,10-11

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z*	document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

19 SEPTEMBER 1996

Date of mailing of the international search report

15 OCT 1996

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

CLAIRE KAUFMAN

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/11452

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- A	MARGOLIS et al. High-efficiency Expression/cloning of epidermal growth factor-receptor-binding proteins with Src homology 2 domains. Proc. Natl. Acad. Sci. October 1992, Vol.89, pages 8894-8898, especially Fig. 3 and "Materials and Methods" section.	1,3,5,7-9 ----- 2,4,6
X --- A	OOI et al. The Cloning of Grb10 reveals a new family of SH2 domain proteins. Oncogene. 1995, Vol.10, pages 1621-1630, especially page 1621, last paragraph, through page 1623, 1st paragraph.	1,3,5,7-9 ----- 2,4,6
X --- A	AUFFRAY et al. IMAGE: Integrated molecular analysis of the human genome and its expression. C.R. Acad. Sci. 1995, Vol.318, pages 263-272, especially abstract and GenBank Accession NO:Z43779.	1,5 ----- 2-4,6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/11452

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/11452

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

- Group I, claim(s) 1-11, drawn to protein (GrbIR-1), polynucleotide, vector, host cell, and method of producing protein.
- Group II, claim(s) 12, drawn to antisense oligonucleotide.
- Group III, claim(s) 13-16, 18, 20, and 22, drawn to modulators of GrbIR-1.
- Group IV, claim(s) 17 and 19, drawn to methods of detecting the presence of modulator.
- Group V, claim(s) 21, drawn to method of treating with modulator.
- Group VI, claim(s) 23, drawn to method of diagnosing GrbIR-1 deficiency.
- Group VII, claim(s) 24, drawn to method of treating with polynucleotide.
- Group VIII, claim(s) 25, drawn to method of treating with GrbIR-1.
- Group IX, claim(s) 26, drawn to transgenic non-human animal.

The inventions listed as Groups I-IX do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Group I-III and IX pertain to products, yet the nucleotide and its encoded protein, the vector, host cell, and cell line of Group I, the antisense oligonucleotide of Group II, the modulator of Group III, and the transgenic animal of Group IX do not share the same or corresponding special technical features of structure and/or function. This includes the polynucleotide of Group I which encodes the protein, while the antisense oligonucleotide of Group II inhibits the protein's expression. The modulators of Group III and transgenic animal of Group IX do not share the same or corresponding special technical features between each other nor between the products of Groups I and II because the modulator must bind the protein or a binding partner and, therefore, is functionally unrelated to the antisense oligonucleotide and the protein itself or DNA encoding it. The transgenic animal is structurally and functionally unrelated by a special technical feature to all other products. The products of Groups I-III and IX are also unrelated functionally to the processes of Groups IV-VIII. The processes of Group I and IV-VIII are performed with materially different process steps and do not share a corresponding special technical feature. Group I is a method of producing a protein and relies on in vitro methods and requires recovery of the protein and, therefore, the functions and processes do not correspond to those of the method of Group VII, which requires isolating cells from a patient and readministration of those cells to the patient. Group IV deals with a method of identifying a modulator and does not share a special technical feature of function with any of the other methods or processes in that the method employs testing binding to GrbIR-1 or its binding partner. Group V is a method of treating a patient with the modulator and does not require GrbIR-1 and, therefore, does not share the same or corresponding technical features. Group VIII is a method of treating a patient with GrbIR-1 and does not require the modulator or a polynucleotide. Group VI is a method of diagnosing GrbIR-1 deficiency by nucleotide sequence comparison and requires process steps and components that do not share the same or corresponding special technical feature of function or structure with those of methods of Groups I, IV-V, and VII-VIII. For these reasons, the respective inventions are not so linked by a special technical feature.

